UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/691,424	10/22/2003	Michael S. Kaplan	M61.12-0988	6179

27366 7590 09/17/2008

WESTMAN CHAMPLIN (MICROSOFT CORPORATION)
SUITE 1400
900 SECOND AVENUE SOUTH
MINNEAPOLIS, MN 55402-3244

EXAMINER	
COLUCCI, MICHAEL C	

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
09/17/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/691,424	KAPLAN ET AL.	
	Examiner	Art Unit	
	MICHAEL C. COLUCCI	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-17* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-17* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

NOTE: Examiner acknowledges the cancellation of claims 18-24.

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 06/13/2008 has been entered.

Response to Arguments

2. Applicant's arguments filed 07/07/08 have been fully considered but they are not persuasive.

Argument 1 (page 6 paragraph 5):

- "Lisle neither teaches nor suggests even the notion of a compression or a compression type. Lisle does discuss collating characters in alphabetical order. However, Lisle neither teaches nor suggests the notion of a compression, in which a plurality of different linguistic symbols are grouped together to form a single sort unit, or a compression table that identifies the compression type (the number of symbols in a given compression)."

Response to argument 1:

Examiner takes the position that Lisle teaches that If a 2 byte compression dictionary or memory is constructed, it will have a possible maximum of 65,536 entries defined. However, one bit must be reserved to indicate that a two-byte pattern is used, so only 32,768 actual dictionary entries are then possible. Viewed differently, 128 segments of 256 words each may be encoded. In many cases, however, the normal vocabulary in constant use for a specific group such as an accountant's or lawyer's office would be only a few thousand words. If the maximum basic dictionary size of 2 bytes or 65,536 possible entries is selected, at least one bit of the first byte would have to be used to identify the fact that a 2 byte encoding pattern is employed because the decompression routine must know how to group the compressed bytes either singly in 2's or in 3's (Lisle Col. 6 lines 25-40).

Further, Lisle teaches a scanning and analysis technique that incorporates counting both the number of characters in each unique word and the number of occurrences of the word within the general usage over a sample of texts from the user's environment. Multiple such dictionaries can be constructed and applied to maximum beneficial effect to achieve a high degree of compression for an individual user (Lisle Col. 4 lines 25-35).

Furthermore, Katayama et al. US 6260051 B1 US 5550541 A (hereinafter Katayama) has been introduced to address the amendment(s) to the claims.

Katayama teaches a registration two-character chain table producing unit 194 for producing a first table block, in which a plurality of registration first and second two-character chains respectively including the same type of for general character and the position numbers of the registration first and second two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore general character type, producing a second table block, in which a plurality of registration special two-character chains respectively including the same type of fore symbolic character and the position numbers of the registration special two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore symbolic character type, and combining each first table block corresponding to one type of fore general character and one second table block corresponding to one type of fore symbolic character determined in correspondence to the type of the fore general character to form a two-character chain table for each character group, the fore characters of the chains in each two-character chain table belonging to the same character group (Katayama Col. 130 lines 33-53).

Further, Katayama teaches a character chain collating and judging unit 200 for receiving the position numbers of one particular two-character chain Tc1 from the storing unit 195 just after the reception of the position numbers of another particular two-character chain Tc2 under the control of the control unit 199. (First collation case), collating each position number of a particular second

two-character chain Tc1 with a particular position number of a particular first two-character chain Tc2 to judge whether or not each position number of the particular second two-character chain Tc1 agrees with the particular position number of the particular first two-character chain Tc2 (second collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular first two-character chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular first two-character chain Tc2 by one (third collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular second two-character chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular second two-character chain Tc2 by two (fourth collation case), collating each position number of a particular first two-character chain Tc1 with a particular position number of a particular special two-character chain Tc2 to judge whether or not each position number of the particular first two-character chain Tc1 is higher than the particular position number of the particular special two-character chain Tc2 by one (fifth collation case), and detecting a particular position number of a particular two-character chain of the particular two-character chain table Tc1 for each collation case (Katayama Col. 131 lines 27-67).

Additionally, Examiner takes Official Notice that it is well known to compress input characters/symbols into a different representation such as the dictionary methods (LZX, LZ78, LZFG, LZRW1, LZRW4, LZW, LZMW, LZAP, LZY,LZ, and other variants of dictionary compression methods taught by "*David Solomon, Data Compression – The Complete Reference, 4th Ed.*, Pages 172-224 & Tables 3.25 and 3.26" demonstrate input and output variants, wherein single symbols/characters are sorted and represented as a combination of symbols/characters.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claim 1, 6, 11, 12, and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lisle et al US 4,843,389 (hereinafter Lisle) in view of Katayama et al. US 6260051 B1 US 5550541 A (hereinafter Katayama).

Re claims 1 and 6, Lisle teaches a computer-readable medium having computer-executable instructions for performing steps for building a symbol table for storing sort weights for a plurality of linguistic symbols used in a plurality of languages supported by a computer system (Col. 15 lines 45-63), comprising:

constructing the symbol table (Col. 19 lines 36-59) to contain a list of code points (Col. 20 lines 35-56) each uniquely identifying one of the symbols, and a sort weight for the symbol identified by said each code point (Col. 15 lines 45-63);

providing a plurality of compression tables (Col. 19 lines 36-59), each compression table pertaining to one of the supported languages (Col. 5 lines 21-50) and having a compression type and containing compressions of symbols of that compression type (Col. 15 lines 45-63);

for each code point in the symbol table (Col. 20 lines 35-56), sorting the compression tables (Col. 19 lines 36-59) to identify a highest compression type our compressions beginning with the symbol (Col. 15 lines 45-63) identified by said each code point (Col. 20 lines 35-56);

storing in the symbol table a tag for each code point to indicate said highest compression type for said each code point (Col. 20 lines 35-56).

NOTE: Tagging a code point is construed to be both functionally equivalent and equally effective as ranking or ordering a code point or address in memory for the purposes of a hierarchical classification.

However, Lisle fails to teach each compression being a grouping of two or more symbols treated as a single unit for purposes of linguistic sorting and the compression type identifying a number of symbols in a given compression

Katayama teaches a registration two-character chain table producing unit 194 for producing a first table block, in which a plurality of registration first and second two-character chains respectively including the same type of for general character and the

position numbers of the registration first and second two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore general character type, producing a second table block, in which a plurality of registration special two-character chains respectively including the same type of fore symbolic character and the position numbers of the registration special two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore symbolic character type, and combining each first table block corresponding to one type of fore general character and one second table block corresponding to one type of fore symbolic character determined in correspondence to the type of the fore general character to form a two-character chain table for each character group, the fore characters of the chains in each two-character chain table belonging to the same character group (Katayama Col. 130 lines 33-53).

Further, Katayama teaches a character chain collating and judging unit 200 for receiving the position numbers of one particular two-character chain Tc1 from the storing unit 195 just after the reception of the position numbers of another particular two-character chain Tc2 under the control of the control unit 199. (First collation case), collating each position number of a particular second two-character chain Tc1 with a particular position number of a particular first two-character chain Tc2 to judge whether or not each position number of the particular second two-character chain Tc1 agrees with the particular position number of the particular first two-character chain Tc2 (second collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular first two-character

chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular first two-character chain Tc2 by one (third collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular second two-character chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular second two-character chain Tc2 by two (fourth collation case), collating each position number of a particular first two-character chain Tc1 with a particular position number of a particular special two-character chain Tc2 to judge whether or not each position number of the particular first two-character chain Tc1 is higher than the particular position number of the particular special two-character chain Tc2 by one (fifth collation case), and detecting a particular position number of a particular two-character chain of the particular two-character chain table Tc1 for each collation case (Katayama Col. 131 lines 27-67).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Lisle to incorporate compression being a grouping of two or more symbols treated as a single unit for purposes of linguistic sorting and the compression type identifying a number of symbols in a given compression as taught by Katayama to allow for control of several letters/symbols within a character chain, wherein a table is used to track, judge, and find the position and number of symbols present (Katayama Col. 130 lines 33-53).

Re claims 11 and 16, Lisle teaches a computer-readable medium having computer-executable instructions for performing steps for a computer search program to carry out a linguistic sorting operation (Col. 15 lines 45-63, comprising:

receiving an input string containing a plurality linguistic symbols (Col. 6 lines 42-58) used in a given language (Col. 15 lines 45-63);

for a first symbol in a combination of symbols in the input string (Col. 15 lines 45-63), referencing a symbol table (Col. 20 lines 35-56) to obtain a highest compression type for compressions beginning with said first symbol (Col. 19 lines 36-59), the symbol table having a list of code points each uniquely identifying a symbol and a sort weight (Col. 15 lines 45-63) for the symbol identified by said each code point (Col. 20 lines 35-56);

performing a binary search (Col. 16 lines 6-27) through each of a plurality of compression tables (Col. 19 lines 36-59) containing compressions for the given language to find a matching compression that matches said combination of symbols in the input string (Col. 16 lines 6-27), wherein the plurality of compression tables are searched in a descending order (Col. 15 lines 45-63) of compression types of the compression tables (Col. 19 lines 36-59) starting with a compression table having a compression type equal to said highest compression type for said first symbol (Col. 15 lines 45-63).

NOTE: Tagging a code point is construed to be both functionally equivalent and equally effective as ranking or ordering a code point or address in memory for the purposes of a hierarchical classification.

However, Lisle fails to teach each compression being a grouping of two or more symbols treated as a single unit for purposes of linguistic sorting and the compression type identifying a number of symbols in a given compression

Katayama teaches a registration two-character chain table producing unit 194 for producing a first table block, in which a plurality of registration first and second two-character chains respectively including the same type of for general character and the position numbers of the registration first and second two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore general character type, producing a second table block, in which a plurality of registration special two-character chains respectively including the same type of fore symbolic character and the position numbers of the registration special two-character chains are listed in the order of arranging the chains in the converted registration character string, for each fore symbolic character type, and combining each first table block corresponding to one type of fore general character and one second table block corresponding to one type of fore symbolic character determined in correspondence to the type of the fore general character to form a two-character chain table for each character group, the fore characters of the chains in each two-character chain table belonging to the same character group (Katayama Col. 130 lines 33-53).

Further, Katayama teaches a character chain collating and judging unit 200 for receiving the position numbers of one particular two-character chain Tc1 from the storing unit 195 just after the reception of the position numbers of another particular two-character chain Tc2 under the control of the control unit 199. (First collation case), collating each position number of a particular second two-character chain Tc1 with a particular position number of a particular first two-character chain Tc2 to judge whether or not each position number of the particular second two-character chain Tc1 agrees with the particular position number of the particular first two-character chain Tc2 (second collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular first two-character chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular first two-character chain Tc2 by one (third collation case), collating each position number of a particular special two-character chain Tc1 with a particular position number of a particular second two-character chain Tc2 to judge whether or not each position number of the particular special two-character chain Tc1 is higher than the particular position number of the particular second two-character chain Tc2 by two (fourth collation case), collating each position number of a particular first two-character chain Tc1 with a particular position number of a particular special two-character chain Tc2 to judge whether or not each position number of the particular first two-character chain Tc1 is higher than the particular position number of the particular special two-character chain Tc2 by one (fifth collation case), and detecting a particular position number of a

particular two-character chain of the particular two-character chain table Tc1 for each collation case (Katayama Col. 131 lines 27-67).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Lisle to incorporate compression being a grouping of two or more symbols treated as a single unit for purposes of linguistic sorting and the compression type identifying a number of symbols in a given compression as taught by Katayama to allow for control of several letters/symbols within a character chain, wherein a table is used to track, judge, and find the position and number of symbols present (Katayama Col. 130 lines 33-53).

Re claim 12, Lisle teaches a computer-readable medium as in claim 11, wherein the compressions in each of the compression tables (Col. 19 lines 36-59) are sorted according to code points for symbols forming the compressions (Col. 15 lines 45-63).

5. Claim 2, 5, 7, 10, 15, and 17, are rejected under 35 U.S.C. 103(a) as being unpatentable over Lisle et al US 4,843,389 (hereinafter Lisle) in view of Katayama US 6260051 B1 (hereinafter Katayama) and further in view of Ho 7,130,470 B1 (hereinafter Ho).

Re claim 2, 7, and 15, Lisle in view of Katayama fails to teach the

Ho teaches that in current multilingual database architectures, Unicode is often used to depict characters. Unicode is a superset of the ASCII character set that uses two bytes for each character rather than one. Because Unicode is able to handle

65,536 character combinations rather than just 256, it can house the alphabets of most of the world's languages. Unicode is a desirable character set because it easily enables a database user to enter in records in a number of different languages. Alphabetic sort orders are performed by determining a character's sort weight by accessing a collation weight table. A collation weight table provides a numerical value for a character for sorting. For example, in a collation weight table the letter "A" may have a sorting weight of 10 and the letter "B" may have a sorting weight. When the sort order performs the sort, "A" will be ordered before "B" because it has a smaller value. The collation weight table can have a sort weight for every Unicode character, thus allowing for sorting in multiple languages (Ho Col. 1 line 52 – Col. 2 line 3).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Lisle in view of Katayama to incorporate a Unicode standard for assigning code points to symbols as taught by Ho to allow for proper collation, decomposition, and manipulation of text for various writing systems (Ho Col. 1 line 52 – Col. 2 line 3).

Re claim 17, Lisle teaches the computer-readable medium as in claim 11, having further computer-executable instructions for storing a sort weight (Col. 15 lines 45-63) for the matching compression (Col. 16 lines 6-27).

Re claims 5 and 10, Lisle teaches the computer-readable medium as in claim 1, further comprising computer-executable instructions for performing steps of sorting

compressions (Col. 15 lines 45-63) in each of the compression tables based on combinations of code points (Col. 20 lines 35-56) of the compressions in said each compression table (Col. 19 lines 36-59).

6. Claims 3, 4, 8, 9, 13, and 14, are rejected under 35 U.S.C. 103(a) as being unpatentable over Lisle et al US 4,843,389 (hereinafter Lisle) in view of Katayama et al. US 5550541 A (hereinafter Katayama) and further in view of Edberg 5,873,111 A (hereinafter Edberg).

Re claims 3, 8, and 14, Lisle teaches sort weight of the symbol (Col. 15 lines 45-63) identified by said each code point (Col. 20 lines 35-56).

However Lisle in view of Katayama fails to teach the computer-readable medium as in claim 1, wherein the tag for each code point is stored as a portion

Edberg teaches character attributes that may be organized in a particular collation order such that information located earlier in the list indicate a higher priority level of significance. For example, if "number" comes before "letter" in the order of the character attributes in class 40, then any number will be collated before any letter, such that "10" will be listed before "apple" in a list of information which has been collated by the sample ordering of category 32a. Alternatively, the character attributes 46 may be tagged with a prefix 43. The lower the prefix 43 of a character attribute 46, the earlier it places in the collation order. For example, in the Unicode category 32c, Latin letters would list before Cyrillic letters in a collation order (Edberg Col. 12 lines 7-12).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention to modify the system of Lisle in view of Katayama to incorporate the tag for each code point stored as a portion as taught by Edberg to allow for proper ordering and collation of characters, wherein prefixes are considered in a language specific text (i.e. Unicode and/or Latin) (Edberg Col. 12 lines 7-12).

Re claims 4 and 9, Lisle teaches computer-readable medium as in claim 3, wherein the sort weight of the symbol identified by said each code point (Col. 20 lines 35-56) comprises a case weight value (Col. 15 lines 45-63), and wherein the tag for said each code point is stored as part of the case weight value for said each code point (Col. 20 lines 35-56).

Re claims 13, Lisle teaches computer-readable medium as in claim 12, wherein each code point in the symbol table includes a tag indicating a highest compression type (Col. 19 lines 36-59) for said each code point (Col. 20 lines 35-56), and wherein said step of referencing retrieves the tag for the code point identifying said first symbol (Col. 15 lines 45-63).

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-

270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626
(571)-270-1847
Michael.Colucci@uspto.gov

/Richemond Dorvil/
Supervisory Patent Examiner, Art Unit 2626